开放获取期刊集成服务权益问题研究*

章岑[1,2]，刘静羽[1]，黄金霞[1]

1 中国科学院文献情报中心 北京 100190，2 中国科学院大学 北京 100049

[摘要][目的/意义]为解决开放资源集成服务、再利用服务提供权益保障分析方法和解决方案。**[方法/过程]**研究开放获取期刊集成平台的集成和服务方式，分析开放获取期刊集成平台建设中遇到的采集、组织、再利用权益纠纷，包括禁止机器自动采集问题、数据库权保护、数据调用的程度受限。**[结果/结论]**分析当前国内外已有的信息资源使用政策、法律条款，设计OA 期刊集成服务权益管理方案，从数据采集前和出版商进行有效沟通，提供服务时原文链接的清楚标明，最大程度提供第三方数据调用等方面展开，为图书馆开展开放资源集成服务建设提供参考与借鉴。

[关键词]开放资源集成服务；开放获取期刊；权益管理

[分类号] G251

1. 引言

自 2002 年《布达佩斯开放获取先导计划》和 2003 年《关于自然与人文科学知识领域的开放获取柏林宣言》发表以来，开放获取运动得到各方普遍关注并在全球飞速发展，持续推动开放资源的开放、共享和使用。目前越来越多的主体参与到开放获取运动，各国政府部门、科研机构、出版商、图书馆纷纷也出台了大量的开放获取政策，规范和推动资源的开放获取。2015 年，加拿大科技部发布开放获取政策，要求所有由三大联合资助机构（NSERC、SSHRC、CIHR）之一资助的同行评议期刊出版物都应在 12 个月内在线可自由获取[1]；美国国家科学基金会（NSF）发布公共获取计划，增强 NSF 资助研究的科学出版物和数字化科研数据的公共获取[2]；学术出版与学术资源联盟（SPARC）呼吁白宫采取行政措施，确保联邦政府资助的教育材料可作为开放教育资源（OER）自由使用、共享和完善[3]。越来越多数量和类型的信息资源走向开放，开放资源的学术影响力和价值持续增强[4]。

开放获取环境下，图书馆的资源建设工作也在发生变化，通过开展开放资源的建设来扩展馆藏。中国科学院文献情报中心从 2009 年开始逐步建立了开放会议采集服务系统、开放社会经济信息采集服务系统、开放课件采集服务系统、OA 期刊采集服务体系（GoOA）等，通过对开放资源的发现、遴选、评价、保存、再利用建设，提供开放资源的一站式发现。但随着开放资源建设和服务的深入，开放资源例如开放获取期刊（OA 期刊）内容的采集、组织、再利用建设遇到了一系列与权益相关的问题，包括出版商阻碍的采集问题，不提供开放接口、数据库权保护、爬虫协议；集成平台建立检索系统的内容存储问题；以及提供全文下载、数据挖掘、数据调用等深度再利用服务中权益问题，有必要对这些权益问题进行深入研

---

*本文系 NSTL 开放资源建设项目“开放资源权益管理规范及其应用研究”。

作者简介：章岑，硕士研究生，ORCID：0000-0002-1956-352X；刘静羽，中科院文献情报中心助理馆员，ORCID：0000-0003-4115-3524；黄金霞，中科院文献情报中心副研究馆员，通讯作者，邮箱 huangjx@mail.las.ac.cn，ORCID：0000-0001-8705-0067。

究，并提出针对解决方案。

目前，国内外对 OA 期刊集成服务权益问题的相关研究集中于两方面。一方面，集中在 OA 期刊集成服务平台自身的研究，如 Stenson L[5]对 DOAJ 等集成平台提供的服务进行分析，Pandita R[6]从 DOAJ 数据中分析 OA 期刊地理分布和发展情况等，Li, Xuemei[7]等通过引文分析 PMC 等集成平台的学术传播作用。但是其中对集成平台的权益问题关注不够，鲜有相关研究。另一方面，体现在开放获取领域的权益问题，包括版权、许可协议、知识产权等，如 Casey, J[8]阐述了知识共享协议的实践，Kim, M[9]阐述了知识共享协议的使用等，叶兰[10]等从图书馆传统教学角度，研究 MOOCs 的版权涉及的利益相关者及版权问题。但同样没有关注开放集成平台建设者会涉及的权益问题。

图书馆已经从简单的目录索引服务增殖为更深层次再利用服务，而已有法律法规及开放获取政策无法完全解决开放资源再利用中的权益约束问题。因此本文将从图书馆角度，以 OA 期刊建设中最关键的集成和服务环节为例，总结 OA 期刊集成服务方式中的权益问题，寻找问题解决的原则、依据，并提出 OA 期刊采集、组织、服务环节中 OA 期刊再利用权益管理方案，设计权益问题发生后采取的有效处理方式，以期为图书馆开展开放资源集成服务建设提供参考与借鉴。

2. OA 期刊集成方式和服务方式分析

2.1 OA 期刊集成平台的集成方式

当前 OA 期刊集成平台众多，主要包括 OA 期刊出版商出版平台、第三方的 OA 期刊集成平台。前者的 OA 期刊出版平台是由 OA 期刊出版商通过出版平台提供资源集成检索和下载，例如 BMC、Hindawi、Medknow 等出版社平台，只整合了自己出版社出版的 OA 期刊；后者 OA 期刊集成平台是由相关机构集成了众多出版社的 OA 期刊及其内容，进行内容的重新组织并提供更多的服务。后者 OA 期刊集成平台的资源集成方式，主要依赖于前者出版平台能提供的资源获取方式。从分析结果来看，这两者之间存在着不匹配程度，后者希望前者能提供更合适的方式。另外，第三方 OA 期刊集成平台又可分为图书馆等信息服务机构建立的 OA 期刊集成平台（例如 GoOA）和其他 OA 期刊集成平台（例如 DOAJ、PMC）。两者虽然都对前者资源进行集成，但是图书馆 OA 期刊集成平台也会利用其他 OA 期刊集成平台发现资源，更侧重用户需求并提供深度服务。所以，这里立足于出版商 OA 期刊集成平台和其他第三方 OA 期刊集成平台提供的资源获取方式分析，重点分析图书馆 OA 期刊资源集成方式。

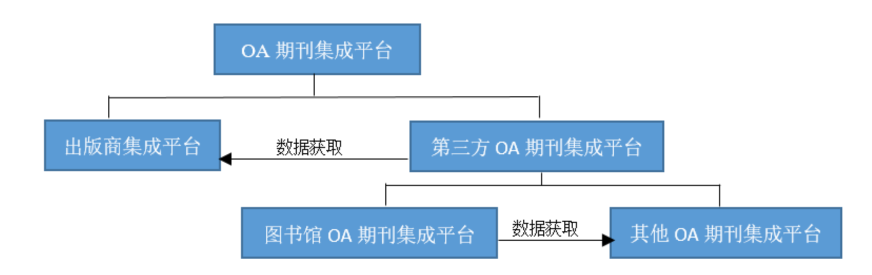


图 1. OA 期刊集成平台类型

2.1.1 出版平台提供的资源获取方式

第一种主要方式：OAI 协议或 FTP 传输。BMC、Hindawi、Springer Open 等出版商提供获取元数据的标准接口，并且支持通过 FTP 或 OAI 进行全文的批量下载，尽量做到最大化开放共享。PLoS 支持自动将论文发布到相关的机构仓储或学科仓储，保证开放论文及时存储到一些常用的开放仓储中（如 PMC）。表 1 列出六大主要出版商的开放政策，其中完全开放获取出版商相对支持 OAI 或 FTP 的方式提供数据，而传统开放获取出版商如 Elsevier 的开放力度较弱，对资源的限制性较强。

出版商	第三方接口	数据调用方式
BMC	OAI-PMH、RESFUL API、RSS	FTP
Elsevier	禁止	禁止
Hindawi	OAI-PMH	FTP
eLife	eLifeAPI	无明确说明
PLOS	Search API、ALM API	存储于 PMC
Springer Open	OAI-PMH、RESFUL API、RSS	FTP

表 1. 出版商数据调用方式

第二种主要获取方式：允许机器爬取。出版商在无法提供 FTP 或 OAI 接口情况下，进行资源采集只能通过机器爬取方式，如搜索引擎获取资源的方式，即通过网页抓取各种格式，将文件中的文本信息提炼出来。虽然没有出版商明确在网站政策中表示支持机器爬取的方式，但是在与出版商沟通协商，存在合作的可能。

2.1.2 OA 期刊集成平台的资源采集方式

PMC 作为数据仓储，通过相关出版商或作者提交资料进行资源整合，存储期刊中的文献和作者手稿，免费向用户提供全文及相关材料。中国科学院文献情报中心 GoOA 也通过与出版商合作协议，目前已有 8 家出版商签署了合作协议，同意提供 FTP 或开放接口方式提供资源，另如 Oxfold、Elife 可以支持 GoOA 自行进行数据爬取，并将 robots 协议进行相对宽泛的设置。

由于 OA 期刊集成平台主要依赖出版商合作获取资源，采集环节主要通过与出版商合作或作者自己递交。采集环节作为 OA 期刊集成平台建设的基础和关键，如果能通过与出版商合作，实现合作提交、签订协议、支持机器爬取等提供资源方式，可以有效解决图书馆在 OA 期刊采集建设中可能产生的纠纷。

2.2 出版商平台和 OA 期刊集成平台的服务方式

OA 期刊集成平台建设目标是为了提供开放资源的服务，服务方式的不同将直接影响到平台的使用。除了提供用户访问、下载 OA 期刊论文的常规服务，因为 OA 期刊本身具有的开放性、版权清晰等特征，使得 OA 期刊集成平台将提供一些特别的服务方式。

2.2.1 集成检索

集成检索服务是指通过标准协议（如 OAI-PMH）收割元数据及获取全文，建立本地索引，将来自多个机构的开放资源集成和整合在一起形成新的使用方式，为用户提供检索和下载服务[11]。例如 OpenAIRE 通过机构仓储，将欧盟范围内的开放获取论文集中起来，提供

集成检索；学术搜索引擎 BASE Search，提供对全球异构学术资源的集成检索服务。集成检索通过提供检索、浏览、下载或链接到原文，是 OA 期刊集成平台最基础和根本的服务方式。

2.2.2 科研数据管理

科研数据管理是期刊可持续发展的重要资源，国外诸多 OA 期刊已经实现数据管理。首先体现在数据政策上，出版商执行强制或非强制性数据存储政策，Nature、Science、PLoS 等强制要求作者发表论文时提交论文数据或者论文的补充材料[12]。其次，数据的关联发现方面。论文发表时把研究数据提交到数据仓储中，通过数据唯一标识符与论文关联[13]。BMC 目前约 50 种期刊已采取在发表的论文中提供数据链接，支持多种存储形式，数据必须永久开放。再次，数据服务方面，PLoS 为作者推荐数据知识库列表、数据质量审查、提供数据 DOI。BMC 通过与 LabArchives 合作，在作者提交手稿后，可以获得来自 LabArchives 的 100M 免费存储服务[14]。

2.2.3 语义标注和数据挖掘

文本和数据挖掘（Text Mining）是指从文本数据中抽取有价值的信息和知识的计算机处理技术[15]，从而获得科学发现。出版商大多鼓励对数据进行挖掘，文本挖掘结果用于非商业性使用。BMC 声明所有数据内容是开放获取，供研究人员进行挖掘。BMC 还提供机器可读的全文 XML 格式，便于机器分析和挖掘。Wiley 表示，遵循 CC BY-NC-ND 和 CC BY-NC 的，可以用于非商业使用进行翻译、文本挖掘。Elsevier 也已经将文本挖掘的权利写入 Science Direct 定购协议中和用户使用协议中，文本挖掘结果可以作为非商业性使用。Springer 与著作权结算中心合作，使得企业生命科学研究人员能够通过出版商很快地、很容易地开展文本和数据挖掘[16]。文本和数据挖掘可以帮助研究人员对大量的信息、识别模式和趋势进行整理，了解文本本身和文本之间的联系，使分析大量信息成为可能。

2.2.4 统计服务

统计服务是指能够追踪到研究成果的查看次数和下载次数，如 PLoS 文献计量(ALM)API，主要用途是分析文献对科研人员的价值，包括单篇文章的使用统计、引用情况、社会网络发布量、博客提及量等。通过对文章的数据统计，分析文章的学术价值和影响力。用户还希望从开放获取期刊获得向机构自动和定期传送机构作者所发表论文的各种当期和累积的统计数据（投稿量、发文量、发文速度、下载量、下载率、引用量、引用率以及社交媒介引用链接数据等）。

2.3 图书馆 OA 期刊集成平台集成和服务特点分析

从以上的分析结果来看，图书馆 OA 期刊采集方式和服务方式存在这样的一些特点：

（1） 公益性集成平台，亟待资源提供更大的使用权限。

OA 期刊集成平台由于涉及到作者、出版商多个主体的利益，很多出版平台对资源的访问、使用等提出了一些政策限制、技术限制，不利于资源再利用，数据获取困难。而图书馆作为信息服务机构，在平台建设和数据获取方面的技术已经比较完善，但由于出版商政策和权益限制，同样导致开放资源建设受到阻碍。

（2） 扩展 OA 期刊服务方式，增强集成平台服务内容。

图书馆 OA 期刊集成平台的目的不仅是建立大而全的集成，而是围绕用户需求、制定深

入服务。除了提供 OA 期刊集成平台普遍提供的 OA 期刊下载、供个人阅读或者本地保存、文章查看和下载次数等服务，图书馆 OA 期刊集成平台还将整合出版商平台的服务，如科研数据管理、提供文本和数据挖掘、允许 OA 期刊数据调用的开放接口、形成 OA 期刊的再传播等。

3. 图书馆 OA 期刊集成服务建设流程中的权益问题分析

根据在 GoOA 开放期刊集成平台的建设中的经历，笔者们总结了图书馆 OA 期刊集成平台建设流程（图 1）。首先通过协议采集或机器爬取方式采集出版商数据，其次将采集到的元数据、全文、附加数据进行组织、存储、索引，建成独立的集成检索系统，最后提供多种深层再利用服务。资源采集和利用服务，是开放资源集成服务流程中的重点环节，从前文的分析来看，有必要在建设初始阶段就分别剖析和解决其中的权益问题。



图 2. OA 期刊集成平台建设流程

3.1 OA 期刊采集中的权益问题

3.1.1 内容提供方禁止机器采集

在采集环节中，遇到的最大阻碍是来自目标资源的“禁止机器采集”限制。目前还没有出版商明确标明支持机器爬取的政策，其中只有 Elsevier 明确规定不允许机器爬取，不能使用爬虫软件进行下载。相反，大部分出版商虽没有明确表示禁止机器爬取，但是拥有相对严格的 robots 协议，并对搜索引擎等机器抓取的时间间隔进行了限制。另如 PLoS 规定，要合理使用 API，但任何行为不应给网站的服务器带来负担。图书馆作为公益性、服务型机构，是否有理由与出版商进行数据采集的沟通合作，获取采集权利。

出版商	Robots.txt 协议
BMC	有。规定了禁止机器抓取的页面；抓取间隔不小于 1 秒
Elsevier	有。规定了禁止机器抓取的页面；抓取时间间隔不小于 10 秒
Hindawi	有。规定了禁止机器抓取的页面；无抓取时间间隔限定。
eLife	有。规定了禁止机器抓取的页面；抓取间隔不小于 10 秒。
PLOS	有。规定了禁止机器抓取的页面；抓取间隔不小于 30 秒
Springer Open	有。规定了禁止机器抓取的页面；抓取间隔不小于 1 秒

表 2. 出版商平台机器采集限制

3.1.2 出版商声称数据库保护

传统意义上，出版商对数据库拥有版权，各国法律也有明确规定，如欧盟《数据库保护指令》、中国《信息网络传播条例》都包含对数据库的保护。

开放环境下，出版商一般认为元数据属于术语众包形式的产物，拥有权无法清晰认定，且不具原创性，因此大部分元数据并不受版权的限制和保护。如 DPLA[17]表示单独的元数据记录免费使用，BMC 明确元数据按照 cc0 协议进行共享[18]。但是由元数据构成对构成的数据库，数据库提供方往往表示拥有版权。除了数据库版权，Elsevier 还规定文章一旦发表，作者需要把发表权让渡给出版商。因此考虑图书馆采集开放资源（OA）全文内容时，是否也会涉及出版商对数据库使用的版权申明及使用限制。

3.1.3 内容的许可协议不规范

目前最普遍的许可是知识共享协议（CC 协议），即作者通过选择使用许可协议维护版权并让渡其他权利。但是 CC 协议的使用并不完全统一。以 DOAJ 收录资源来看，截至 2016 年 3 月 15 日，大部分资源遵循 CC 协议，小部分遵循如出版商自行制定协议、无协议或非 CC 形式的协议。许可协议的不统一、不规范，将导致再利用服务中使用混乱，使用规则不一致，容易产生误用。此外，在既包括传统出版模式又支持开放获取的复合期刊中，也可能出现同一本期刊中不同的论文使用许可不一致。

协议种类	DOAJ 收录的非 CC 协议期刊数	DOAJ 收录的非 CC 协议文章数
Not CC-like	43	2,793
BY-NC-ND	无数据	1,281
Publisher’s own license	76	1,280
Publisher’s license agreement	2	无数据
none	12	无数据
Total	133	5,354

表 3. DOAJ 收录的非 CC 协议资源

Elsevier 开放仓储（Open Archive）使用 Elsevier 自行制定的用户使用许可，共 98 本期刊，其中包括 cell 等著名期刊，开放需要 6-48 个月的禁运期。Elsevier 金色开放获取期刊使用 CC 许可，并对 CC 许可和 Elsevier 用户许可的区别进行表 4 说明[19]。主要区别在于再利用方面，Elsevier 用户许可更加严格，不允许再利用。许可协议的不一致，对图书馆 OA 期刊的采集和组织也造成了一定的困扰和障碍。

用户许可	阅读、打印和下载	再分配或再出版（如存储于知识库）	数据和文本挖掘	翻译	再利用或提炼文章中部分	商业用途买卖和再利用
商业 CC BY 4.0	是	是	是	是	是	是
CC BY NC ND 4.0	是	是	是	是（仅供个人）	是	否
Elsevier 用户许可	是	否	是	是	否	否

表 4. Elsevier 使用许可说明

3.2 OA 期刊服务中的权益问题

3.2.1 已采集资源的集成问题

开放集成平台需要将采集获得的 OA 期刊数据，存放于本地服务器，建立本地存储。通过聚合、集成和整理，涉及元数据的再利用、内容的再利用、研究数据或附加数据的再利用等。需要考虑图书馆是否可以进行内容存储，建立开放仓储；图表数据是否可以通过与来源链接进行存储，建立数据仓储，并提供数据的下载。

3.2.2 数据调用程度的限制

开放集成平台中 DOAJ、SCIELO、PMC、GoOA 均提供开放接口的方式，支持第三方数据调用。其中 PMC 作为数据仓储，支持 FTP 传输，但是禁止机器自动采集，网站有 robots 协议，限制机器抓取。因此，图书馆 OA 期刊平台建设中，需要考虑对于第三方用户如机构或搜索引擎的数据抓取要求，OA 期刊集成系统应提供怎样服务，即是否可以将获取到的资源进行组织、存储和再传播、对于用户的行为目的是否加以区分、调用的内容是全文还是元数据、是否提供开放接口等、应该制定相应的网站使用政策和数据使用协议。

集成平台名称	第三方接口	数据调用方式
DOAJ	OAI-PMH、 Bulk OAI、Search API、CRUD API	无明确说明
SCIELO	OAI-PMH	部分期刊存储于 PMC
PMC	OAI-PMH、OA web service	FTP 提供 XML，images, PDF，附加数据文件，禁止机器自动爬取
GoOA	OAI-PMH	支持 FTP 传输，禁止机器自动爬取

表 5.开放集成平台数据调用方式

综上，图书馆 OA 期刊再利用面临的关键问题包括：①如何避免机器抓取数据时与出版商产生纠纷；②开放资源本身存在的许可协议问题；③数字资源的存储整合；④图书馆是否可以对 OA 期刊进行文本挖掘等深度利用；⑤如何应对搜索引擎对图书馆自建 OA 期刊集成系统的数据抓取。

4 图书馆 OA 期刊采集和服务中权益问题管理方案设计

为解决上述权益问题，引入危机管理思想作为图书馆 OA 期刊集成服务权益管理方案设计的基本原则。罗伯特[20]提出 4R 理论，将危机管理分为缩减、预备、反应、恢复四个阶段，即减少危机情境的攻击力和影响力，使企业做好处理危机情况的准备，尽力应对已发生的危机，以及从中恢复。刘兹恒等[21]提出图书馆危机管理内容模型，包括图书馆日常危机管理、图书馆危机时间管理、图书馆危机后续管理等。

因此将图书馆危机管理思想归纳为：事前预防、事中沟通处理和事后突发应急三个方面。图书馆 OA 期刊建设中，首先要尽量避免权益纠纷，其次是当权益问题不可避免发生时，图书馆积极做好应急沟通，最后使用相关法律、政策规定作为协商依据。

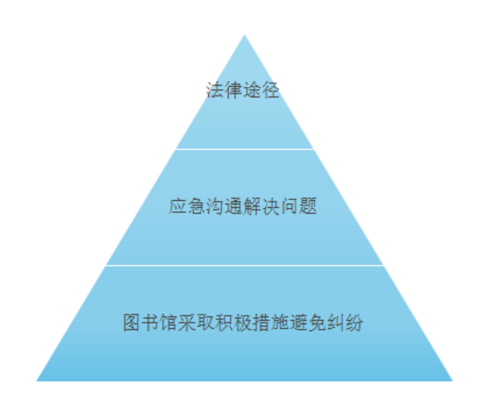


图 3. 图书馆 OA 期刊权益管理方案主要原则

4.1 图书馆采取积极措施避免纠纷

① 采集前与出版商进行积极沟通

在未征求出版商同意的情况下，自行采集开放资源有可能造成对出版商数据库相关权益的侵害，产生纠纷、造成不良后果。因此数据采集前，应积极与内容提供方进行谈判、合作、沟通，争取通过协议、授权的方式，或机器爬取等批量采集方式获取资源，同时需要遵循出版商政策规定和 robots 协议。

② 原文链接需清楚标明

图书馆可以为本地存储数据建立索引、整合提供集成检索和浏览服务或集成系统，提供期刊论文的浏览和下载。但是图书馆需明确标明作品的初始来源（URL），保证作品从图书馆的集成系统可以链接到原网站。定期对平台中资源的使用状况进行统计，并将统计报告提供给出版商，以获取得出版商的更多支持和合作。图书馆必须尽可能保障作品的完整尊重作品的版权和协议、保证作品完整性、标明作品原始来源。

③ 保证数据最大程度传播

图书馆的 OA 期刊服务系统，应保证开放获取资源得到最大方式传播利用。建议提供开放接口供第三方出于任何合理目的使用，或者通过 FTP 向非营利机构批量提供数据；图书馆网站也应允许搜索引擎等抓取图书馆集成系统中的数据。同时作为公共服务机构，有义务指导用户合理使用开放资源，制定集成平台使用政策和用户使用许可。使用许可按作者的原始许可，且文中数据若无单独标明使用协议，默认与原文相同许可。

目前 GoOA 平台已应用到一些实践，包括与出版商制定数据合作协议，采集工作得到多家开放出版商支持；GoOA 对期刊论文的元数据信息予以详细著录，并提供原始网址链接；GoOA 拟定数据使用协议和用户使用协议等。

4.2 应急处理和纠纷解决

如果内容提供方对已采集的数据存有疑义，图书馆应证明图书馆已经采取积极措施，如根据 robots 协议，主动对采集频率进行限制；声明图书馆的公益性，并强调图书馆严格按照期刊论文的使用协议进行下载、使用，标明原文来源，并已经及时停止采集行为。根据 CC4.0 协议，对已采集内容及时进行修改或删除。尽管图书馆作为公益机构，一旦发生侵权行为，也需要承担相应的赔偿责任。

对于始终无法联系到的出版商，数据采集工作又必须进行的情况下，图书馆应如实告知

采集工作情况，争取他们的同意；同时还建议将所有的沟通过程完整记录，并进行公证保全，这些可作为日后处理权益纠纷的有效证据。

4.3 国内外信息资源利用权益的法律和规范分析

在纠纷问题无法沟通解决时，对国内外信息资源利用权益法律规范的研究将为问题的解决提供思路。这里的规范，既包括来自开放获取运动本身的推动目标，也包括一些特有法律或适用于图书馆等公益机构的法律或规则。通过法律和规范的分析，为图书馆开放资源的采集和服务中的权益问题解决提供借鉴。

① "开放获取"本身具有的特别含义

开放获取，从根本上规定了开放获取再利用的权利范围，包括：①开放获取的论文进入互联网公共领域，并且可以被免费获取；②允许任何互联网的用户阅读、下载、复制、传递、打印、搜索和超链接。③任何人都可为资源建立索引，用作软件输入数据或其它任何合法用途[22]。

② 合理使用原则

合理使用是指在某些情况下不经权利人许可或不用支付相应费用、依然有权使用受著作权保护作品的权利，尤其是当作品的使用对于文化或社会来说带来重要益处时。目前国外以"三步检验法"判定著作权合理或侵权，即在某种特殊的情况下、不与作品的正常使用相冲、不致无故损害权利人的合法权益[23]。《美国著作权法案》第 107 条对于合理使用做出了原则性规定，适用于任何立法没有明确规定图书馆相关的合理使用行为。判定"合理使用"的四大要素：使用的目的和性质、享有版权的作品的性质、与享有版权的作品整体相比所使用部分的数量和实质性、使用对享有版权的作品的潜在市场或价值的影响[24]。

譬如 2008 年，剑桥大学出版社、牛津大学出版社和 SAGE 出版公司共同起诉乔治亚州立大学的著作权侵权。出版商认为，该大学在并未获得许可的情况下，使用电子资源储备系统中的版权保护材料，此举违反了版权法。2014 年美国第十一巡回上诉法院就乔治亚州立大学电子藏书系统（e-reserves）案件中发布了判决，拒绝了出版商撤销 e-reserve 服务的尝试，判定合理使用是一个灵活的原则[25]。

合理使用要具体案例具体分析。图书馆作为公益性的文献信息收集、加工、保存和传播机构，增殖开放获取信息传播，以科研传播为目的，不涉及商业化使用，可以利用合理使用制度履行使命。如果在版权保护中对图书馆过度限制，用户将失去获得信息的方便快捷，难以实现信息资源的共享。

③ 著作权文本和数据挖掘例外

目前越来越多的法律、政策向支持文本挖掘的方向发展。英国版权法修改法案增加了文本和数据挖掘例外，具体而言，任何人都可以在无须征得版权人同意的情况下出于非商业目的而对其所读到的材料进行文本和数据挖掘。欧洲版权法也正在努力完善支持数字和文本挖掘的版权例外。美国许多法庭支持再利用，认为将文本挖掘作为合理使用有其必要性，尽管复制数据库的内容是受到著作权保护的。美国虽然没有明确允许文本挖掘具体的限制或例外，但是文本挖掘作为新的研究工具产生和发展被认为属于合理使用[26]。

2011 年，多国作家及作家协会对 HathiTrust、密歇根大学等提起侵权诉，HathiTrust 通

过谷歌获取数字文件引发争议，作家试图阻止未经授权的 700 万本版权保护期内图书扫描版本[27]。2015 年，法院明确支持 HathiTrust 文本挖掘是合理使用，认为 HathiTrust 包含在数据库中的数字化作品允许数据挖掘和文本分析，这使得识别和定位信息来源更容易。

2015 年 12 月，超过 200 个组织和近 500 人签署了《数字时代知识发现海牙宣言》[28]，要求改变知识产权法并且获取技术，认为包括版权法在内的知识产权体系并没有起到支撑开放研究和挖掘内容的作用。呼吁明确知识产权法律的范围，同时为实现内容挖掘呼吁建设更好的基础设施。

④ 避风港原则

“避风港”原则是指当网络服务商实际上不知道也没有意识到能明显推出侵权行为的事实或情况时，在接到权利人的合格通知后，及时移除侵权内容的，不承担责任[29]。也就是说，发生著作权侵权案件时，因行为人未意识到侵权，因而只承担断开链接即停止侵权的责任，不承担赔偿责任。

2011 年，百度文库曾以“避风港原则”为由，认为采用资源分享方式，不必对用户上传的任何作品承担法律责任，也不需要履行合理的版权注意义务。但是百度文库对搜索和链接结果进行了编辑、修改和选择，根据相关法律规定，网站一旦采取了这些行为，就不能适用于“避风港原则”，所以百度文库的行为属于侵权行为。

因此，图书馆适用避风港原则的前提是图书馆不得侵犯版权人的信息网络传播权或者删除、改变版权人的权利管理电子信息[30]。图书馆不能将避风港原则视为“保险箱”，需要在尊重出版商对作品的发表权的情况下使用该条例。“避风港原则”适用于使用程度不明确的资源，图书馆进行数据采集后，发生矛盾可避免与数据商的矛盾激化。“避风港”原则应当可以为图书馆提供免责的依据或抗辩理由。

⑤ CC4.030 天侵权恢复有效期

CC 协议 4.0 是在结合各国版权法和国际知识产权惯例基础上，对以往版本进行的改进。版本的协议规定，一旦侵权行为发生，CC 协议则自动终止；如果还想继续使用资源，必须在 30 天内纠正相关行为。4.0 以前的版本，若使用者违反了协议，则使用者的权利将被立即终止[31]。如果发生侵权行为，只要侵权方立刻改正违反，就能够恢复 CC 协议，而不会永久丧失权利。

综合上述权益管理方案，值得注意的是当前法律对图书馆开放信息环境的尚不明确，图书馆开放资源建设中涉及的权益与著作权规定的状况不尽相同，国外著作权法也正在积极改革。因此图书馆在解决权益问题时，做好权利主体间的沟通是重要前提，尽量在双方均衡的情况下合理进行解决。

5 结语

开放环境已经改变了我们在教育和研究中对知识资源的获取和使用方式，培养出创造力和创新能力，数字时代的图书馆必须保证这些目标的可持续发展和实现。从 DOAJ Seal 的颁布[32]，到 PLOS、SPARC、OASPA 联合发布的开放获取频谱 2.0 版本[33]，再到 NISO 建议增加<free_to_read>和<license_reference>两个元数据字段[34]，可以看出开放资源再利用的程度正逐渐提高，权益信息的机读性也在逐渐规范。我们希望随着开放获取运动的不断深化和开

放基础设施的不断加强，开放资源的再利用方式得以完善，权益纠纷得以充分缓解。

作者贡献说明：章岑，搜集资料，负责撰写文章第 1,2,3 部分，整体统稿及修订论文；
刘静羽，搜集资料，负责撰写文章 4,5 部分；
黄金霞，负责制定文章大纲，提出修改意见。

6 参考文献

[1]Minister holder announces new open access policy for research[EB/OL].[2016-03-25].http://www.science.gc.ca/default.asp?lang=En&n=415B5097-1.

[2]Public access to results of nsf-funded research[EB/OL].[2016-03-25].http://www.nsf.gov/news/special_reports/public_access/.

[3]Sparc calls on white house to open up access to federally funded educational resources [EB/OL].[2016-03-25].http://sparcopen.org/news/2015/sparc-calls-on-white-house-to-open-up-access-to-federally-funded-educational-resources/.

[4]黄金霞,张建勇,黄永文,等.开放资源建设的措施及工作策略[J].2013.57(8):57-61.

[5] STENSON L.Why all these directories?an introduction to doaj and doab[J].Insights,2012,25(3).

[6] PANDITA R. Growing trend towards open access publishing at global level: An analysis of directory of open access journals (doaj)[J]. International research: journal of library and information science,2013,3(3).

[7] Li X, THELWALL M, KOUSHA K. The role of arxiv, repec, ssrn and pmc in formal scholarly communication[J].Aslib journal of information management,2015,67(6):614-635.

[8] CASEY, J. Creative commons licences:are they right for you?Art libraries journal[J],2012 (37/2): 32-37.

[9] KIM M. Uses of creative commons licenses[J].Open source business resource,2008(April 2008).

[10]叶兰,易晓娥.图书馆视角下的 MOOCs 版权问题研究[J].大学图书馆学报,2014(5):25-33.

[11]黄永文,张建勇,谢靖,等.开放资源的再利用模式研究[J].图书情报工作,2013,57(21):32-37.

[12]陈秀娟,吴鸣.学科领域科研数据知识库调研与分析——以化学领域为例[J].图书情报工作,2015(9):111-118.

[13]陈秀娟,吴鸣等.嵌入科研工作流的图书馆数据管理服务——以化学学科为例[J].图书馆论坛,2016(3):1-9.

[14]吴蓉,顾立平等.国外学术期刊数据政策的调研与分析[J].图书情报工作,2015,59(7):99-105.

[15]文本数据挖掘 [EB/OL].[2016-03-25].http://baike.baidu.com/link?url=MZgoiFK2HdagjceXv2mOOywOlbecpvVtv8EG106M_2F7fgJxCtrROw52fL934MOSwTo_kanynBfC9eGGKAfjY_.

[16]Springer partners with copyright clearance center to power text and data mining solution[EB/OL].[2016-03-25].http://www.copyright.com/springer-partners-with-copyright-clearance-center-to-power-text-and-data-mining-solution/.

[17]The digital public library of america policy statement on metadata[EB/OL].[2016-03-25].http://dp.la/info/wp-content/uploads/2013/04/DPLAMetadataPolicy.pdf.

[18]BMC open data policy[EB/OL].[2016-03-25].http://www.biomedcentral.com/about/policies/open-data.

[19]Elsevier user license[EB/OL].[2016-03-25].https://www.elsevier.com/about/company-information/policies/open-access-licenses/elsevier-user-license.

[20]Robert Heath.王成,宋炳辉,金瑛译.危机管理[M].北京:中信出版社,2001:32.
[21]刘兹恒.图书馆危机管理基础工作策略[J].图书馆论坛,2008,28(6):21-25,161.

[22]Budapest Open Access Initiative[EB/OL].[2015-09-09].http://www.budapestopenaccessinitiative.org/read.
[23]Lyman R, Stanley F. Birch. A Unified Theory of Copyright[M].Houston: University of Houston Law Center,2009:355-356.
[24]周玲玲,杨萌等.网络时代图书馆相关公共利益的维护研究——以我国《著作权法》及第三次修法为例[J].图书情报工作,2016,59(23):5-10.
[25]Fair use decision making post-georgia state[EB/OL].[2016-03-25].http://www.arl.org/storage/documents/publications/gsu-code-best-practices-6nov14.pdf.
[26]Text and data mining and fair use in the United States[EB/OL].[2016-03-25].http://www.arl.org/storage/documents/TDM-5JUNE2015.pdf.
[27]多国作家及作家协会对 HathiTrust、密歇根大学等提起侵权诉讼[EB/OL].[2016-03-25].http://www.prnasia.com/pr/2011/09/15/110907521.shtml.
[28]A copyright exception for text and data mining[EB/OL].[2016-03-25].http://libereurope.eu/wp-content/uploads/2015/11/TDM-Copyright-Exception.pdf.
[29]BROWN B. Fortifying the safe harbors: reevaluating the dmca in a web 2.0 world[J]. Berkeley technology law journal, 2008, 23(1): 437-467.
[30]汪琼,陈伟. "避风港" 原则之图书馆适用研究[J].现代情报,2011,31(8):130-133.
[31]新一代的授权方式[EB/OL].[2016-03-25].http://ndaip.sinica.edu.tw/content.jsp?option_id=2701&index_info_id=8644.
[32]DOAJ Seal is now live on the site[EB/OL].[2016-03-25].https://doajournals.wordpress.com/2015/06/11/doaj-seal-is-now-live-on-the-site/.
[33]"How Open Is It?" open access spectrum[EB/OL].[2016-03-25].https://www.plos.org/open-access/howopenisit/.
[34]Access license and indicators[EB/OL].[2016-03-25].http://www.niso.org/apps/group_public/download.php/14226/rp-22-2015_ALI.pdf.

Right Problems about Integration and Service of Open Access Journals

[Abstract] [Purpose/Importance]Providing analysis and methods to solve rights problems in the progress of open access platform about reuse development.**[Methods/Progress]**Discuss several kinds of open access systems about their integration and service, and analyze the rights problems about open access development, problems including rejects machines downloading resourcesautomotive, database protecting problems, and the limits of resources reuse application progress interface. **[Results/Conclusion]**According to policies and terms, design a right management plan, such as communicating with information providers, indicating the links of original article and offering anyone who needs data with noncommercial purpose. In order toprovide suggestions and advice for libraries to develop open access systems.

[Key words] Open access integration and service; Open access journal; Resource reuse management